

# Bayesian Classifier

## I. Motivation

In this experiment we try to classify three sets of data with five different classification models, namely:

- 1) Bayes with  $C$  same for all classes
- 2) Bayes with  $C$  different for all classes
- 3) Naive Bayes with  $C = \sigma^2 I$
- 4) Naive Bayes with the same  $C = \text{diag}(\sigma_1, \sigma_2)$  for all classes
- 5) Naive Bayes with different  $C$  for all classes

## II. Linearly Separable data

Fig1-3 give the PDF ,decision boundary, density curves of one of the models on linearly separable data.

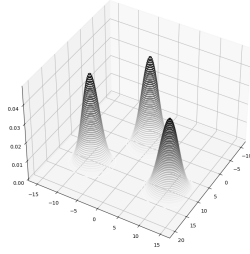


Fig. 1. *PDF for model – 3*

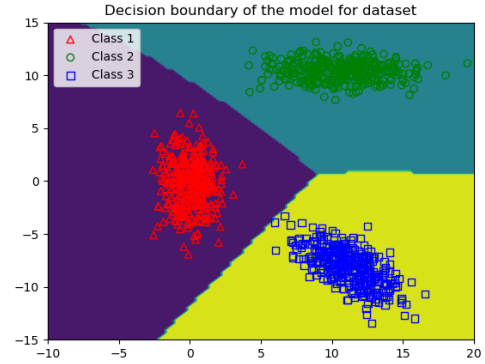


Fig. 2. *Decision boundary for model – 3*

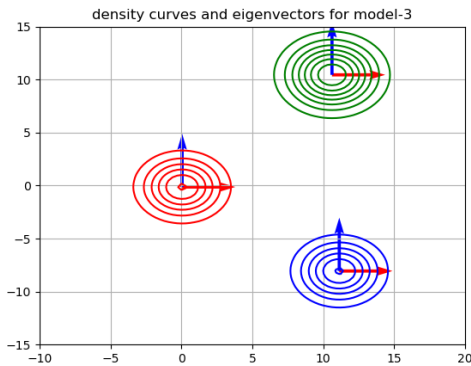


Fig. 3. *Density curves for model – 3*

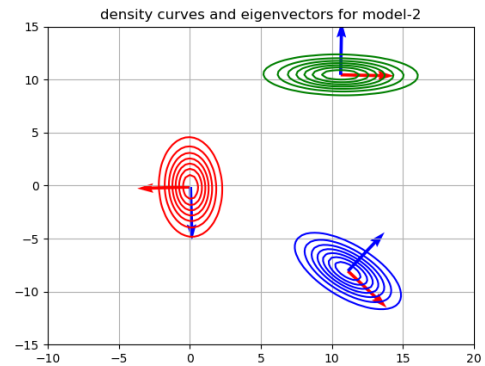


Fig. 4. *Density curves for model – 2*

From fig 3 and 4 we can observe that the models primarily differ in terms of how their Gaussian contours spawn out on the x-y plane. However in this case the clusters(for each class) are wide apart, so the different contours did not have much effect and further 100 percent accuracy is obtained on the development data for all the five models. This is also evident from a perfect ROC curve in fig5.

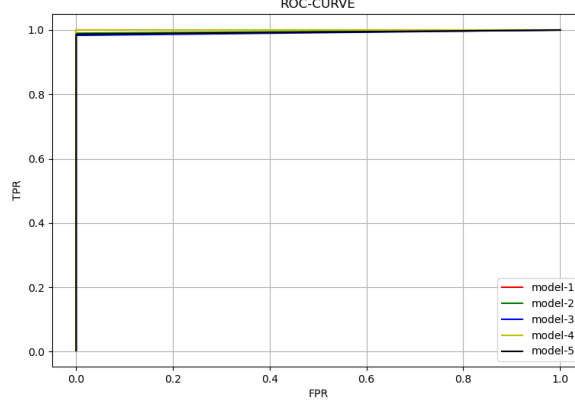


Fig. 5. ROC Curve

### III. Non-linearly Separable data

In this case the clusters are distributed as a set of concentric bands.

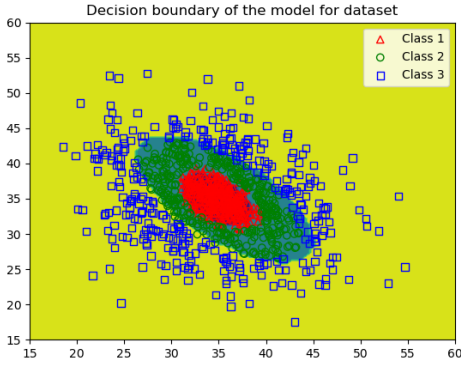


Fig. 6. Decision boundary for model – 2

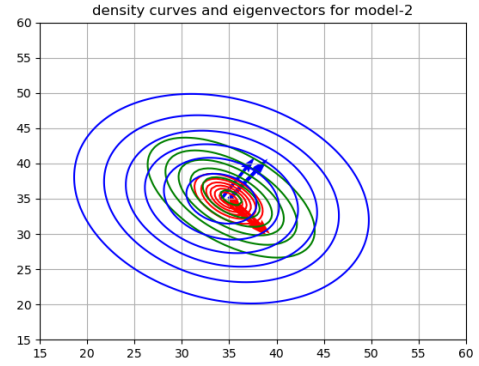


Fig. 7. Density curves for model – 2

Model-2(C different for all classes) appears to perform the best on this data which is evident from its accuracy on development data and the ROC,DET curves.

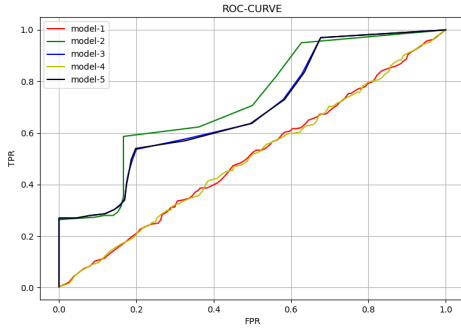


Fig. 8. ROC Curve

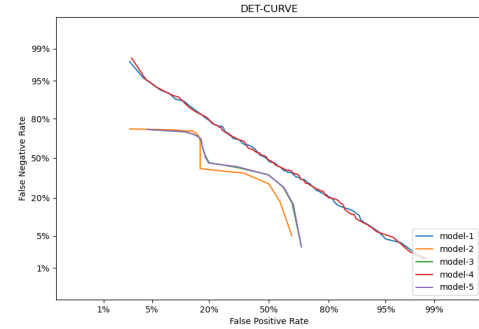


Fig. 9. DET Curve

The Naive bayes classifier have lower accuracy because their contour's eigenvectors are parallel to x and y axes. While from fig-6 it's clear that such contour may not be very helpful, we need a slightly tilted ellipse. Table-1 displays the accuracies obtained on development data for all the models.

Model	Accuracy
1	38.33
2	91.66
3	81
4	38.66
5	81

TABLE I  
ACCURACY FOR NON-LINEARLY SEPARABLE DATA

Model	Accuracy
1	87.33
2	90.33
3	91.33
4	87
5	90.66

TABLE II  
ACCURACY FOR REAL DATA

#### IV. Real Data

For the real data all models seem quite optimal with their accuracies in the ballpark of 87-91. Of which model-3 seems the best, evident from ROC and DET curves too.

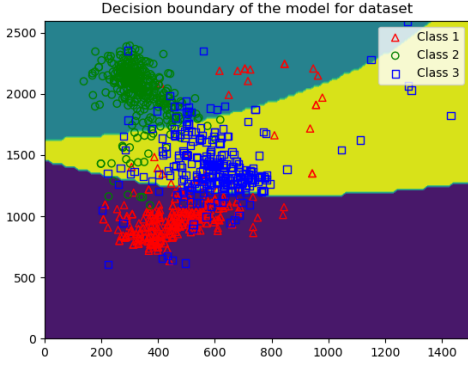


Fig. 10. Decision boundary for model – 3

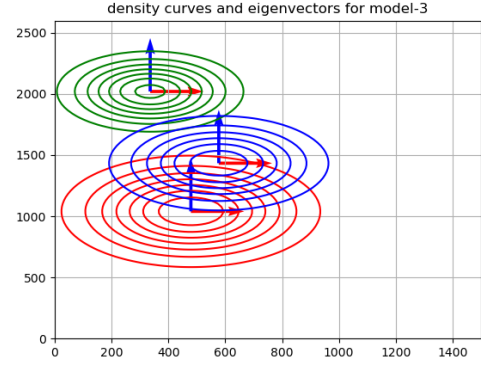


Fig. 11. Density curves for model – 3

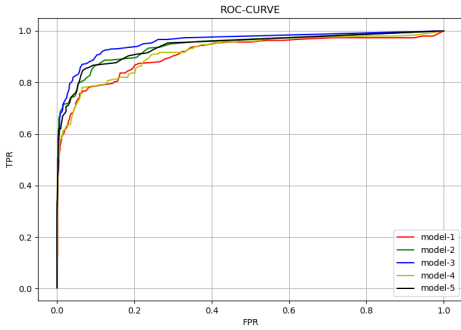


Fig. 12. ROC Curve

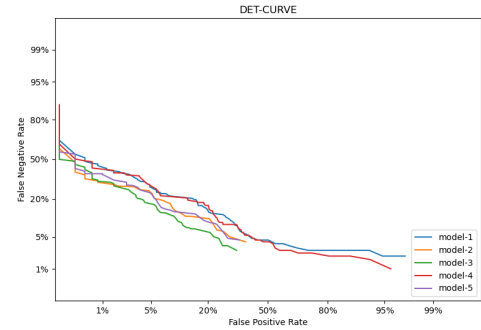


Fig. 13. DET Curve